

Reply to “Codon Usage Frequency of RNA Virus Genomes from High-Temperature Acidic-Environment Metagenomes”

Mark Young,^{a,b,c} Benjamin Bolduc,^{a,d} Daniel P. Shaughnessy,^{a,c} Francisco F. Roberto,^f Yuri I. Wolf,^e Eugene V. Koonin^e

Thermal Biology Institute^a and Departments of Microbiology,^b Plant Sciences and Plant Pathology,^c and Chemistry and Biochemistry,^d Montana State University, Bozeman, Montana, USA; National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA^e; Idaho National Laboratory, Idaho Falls, Idaho, USA^f

We welcome the attention of Stedman et al. to the novel RNA viruses that we recently identified in an archaea-dominated community (1) and their analysis of viral and cellular codon usage as an approach to assist in assigning potential hosts. However, we are not convinced that their analysis of such a limited number of RNA viruses leads to any compelling conclusions without a comprehensive study that should include proper statistical tests.

There are numerous examples of virus codon usage either matching (2–4) or significantly deviating from (5–8) their host cell codon usage. Analyses of codon bias of poxviruses, adenoviruses, enteroviruses, and herpesviruses found a mixture of agreement with and antagonism to the codon bias of their human hosts (9–12). For example, human enterovirus 71 codon usage paradoxically matches the disfavored codon usage of the human host of this virus (11). An analysis of 16 phages of *Staphylococcus aureus* found few correlations in codon usage between a phage and its specific host (13). Indeed, the large variance observed in viral codon usage compared to that of the host cells is likely to be why Stedman et al. found that the *Escherichia coli* bacteriophage Q β can be assigned with either eukaryotic or bacterial hosts. We are unaware of a comprehensive analysis of the host-virus combinations from all three domains of life; we believe that this would be a worthwhile endeavor.

There are many factors that contribute to viral codon usage independent of selection pressure to match the codon usage of its host cell. In some cases, it has been found that genome-wide mutational pressure overrides the selection for specific codons (14); translational selection (15) and protein secondary structure (16) can all contribute. In other cases, codon usage has been linked to the thermophilic lifestyle (17, 18). For example, the codon usage of thermophilic bacteria more closely resembles that of thermophilic archaea than that of mesophilic bacteria (17, 19).

We maintain that the most parsimonious conclusion from the data reported in Bolduc et al. (1) is that the RNA viral genomes we detected replicate in archaeal hosts residing in the sampled Yellowstone hot springs. Additional data compatible with this conclusion include recent single-cell rRNA gene analysis of the originally sampled hot spring showing that 97% of the cells analyzed (68 of 71 cells) represented members of 5 archaeal genera, with 3% belonging to a single bacterial genus. No eukaryotic cells were detected. These results support our original cellular metagenomic analysis reported in Bolduc et al. (1). In addition, we have developed reverse transcription-PCR (RT-PCR) assays for a number of the suspected archaeal RNA viral contigs and can find the same RNA genomes in the viral fraction of the same hot spring 4 years after the original observations as well as in a number of other archaea-dominated Yellowstone hot springs. These findings suggest that the discovered RNA viral genomes do not come from

viruses entering the hot springs from outside sources. Currently, we are actively pursuing direct isolation of the virus together with its host.

REFERENCES

1. Bolduc B, Shaughnessy DP, Wolf YI, Koonin EV, Roberto FF, Young M. 2012. Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. *J. Virol.* 86:5562–5573.
2. Bahir I, Fromer M, Prat Y, Linial M. 2009. Viral adaptation to host: a proteome-based analysis of codon usage and amino acid preferences. *Mol. Syst. Biol.* 5:311.
3. Carbone A. 2008. Codon bias is a major factor explaining phage evolution in translationally biased hosts. *J. Mol. Evol.* 66:210–223.
4. Su M-W, Lin H-M, Yuan HS, Chu W-C. 2009. Categorizing host-dependent RNA viruses by principal component analysis of their codon usage preferences. *J. Comput. Biol.* 16:1539–1547.
5. Gu W, Zhou T, Ma J, Sun X, Lu Z. 2004. Analysis of synonymous codon usage in SARS Coronavirus and other viruses in the Nidovirales. *Virus Res.* 101:155–161.
6. Sau K, Gupta SK, Sau S, Mandal SC, Ghosh TC. 2007. Studies on synonymous codon and amino acid usage biases in the broad-host range bacteriophage KVP40. *J. Microbiol.* 45:58–63.
7. Sau K, Sau S, Mandal SC, Ghosh TC. 2005. Factors influencing the synonymous codon and amino acid usage bias in AT-rich *Pseudomonas aeruginosa* phage PhiKZ. *Acta Biochim. Biophys. Sin. (Shanghai)* 37:625–633.
8. Zhao S, Zhang Q, Liu X, Wang X, Zhang H, Wu Y, Jiang F. 2008. Analysis of synonymous codon usage in 11 human bocavirus isolates. *Biosystems* 92:207–214.
9. Barrett JW, Sun YM, Nazarian SH, Belsito TA, Brunetti CR, McFadden G. 2006. Optimization of codon usage of poxvirus genes allows for improved transient expression in mammalian cells. *Virus Genes* 33:15–26.
10. Das S, Paul S, Dutta C. 2006. Synonymous codon usage in adenoviruses: influence of mutation, selection and protein hydrophathy. *Virus Res.* 117:227–236.
11. Liu YS, Zhou JH, Chen HT, Ma LN, Pejsak Z, Ding YZ, Zhang J. 2011. The characteristics of the synonymous codon usage in enterovirus 71 virus and the effects of host on the virus in codon usage pattern. *Infect. Genet. Evol.* 11:1168–1173.
12. Roychoudhury S, Mukherjee D. 2010. A detailed comparative analysis on the overall codon usage pattern in herpesviruses. *Virus Res.* 148:31–43.
13. Sau K, Gupta SK, Sau S, Ghosh TC. 2005. Synonymous codon usage bias in 16 *Staphylococcus aureus* phages: implication in phage therapy. *Virus Res.* 113:123–131.
14. Jenkins GM, Holmes EC. 2003. The extent of codon usage bias in human RNA viruses and its evolutionary origin. *Virus Res.* 92:1–7.
15. Ghosh TC, Gupta SK, Majumdar S. 2000. Studies on codon usage in *Entamoeba histolytica*. *Int. J. Parasitol.* 30:715–722.
16. Gupta SK, Majumdar S, Bhattacharya TK, Ghosh TC. 2000. Studies

Address correspondence to Mark Young, myoung@montana.edu.

This is a response to a letter by Stedman et al. (doi:10.1128/JVI.02610-12).

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JVI.02883-12

- on the relationships between the synonymous codon usage and protein secondary structural units. *Biochem. Biophys. Res. Commun.* **269**: 692–696.
17. Lobry JR, Necsulea A. 2006. Synonymous codon usage and its potential link with optimal growth temperature in prokaryotes. *Gene* **385**: 128–136.
18. Lynn DJ, Singer GAC, Hickey DA. 2002. Synonymous codon usage is subject to selection in thermophilic bacteria. *Nucleic Acids Res.* **30**:4272–4277.
19. Singer GAC, Hickey DA. 2003. Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. *Gene* **317**:39–47.